



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/805,805	03/22/2004	Mark S. Manasse	MSFT-5031/307238.01	9958

41505	7590	01/15/2008
WOODCOCK WASHBURN LLP (MICROSOFT CORPORATION)		
CIRA CENTRE, 12TH FLOOR		
2929 ARCH STREET		
PHILADELPHIA, PA 19104-2891		

EXAMINER	
STACE, BRENT S	

ART UNIT	PAPER NUMBER
2161	

MAIL DATE	DELIVERY MODE
01/15/2008	PAPER

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

## Office Action Summary

Application No.

10/805,805

Applicant(s)

MANASSE, MARK S.

Examiner

Brent S. Stace

Art Unit

2161

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

### Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

### Status

- 1) ☒ Responsive to communication(s) filed on 13 November 2007.
- 2a) ☒ This action is **FINAL**. 2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

### Disposition of Claims

- 4) ☒ Claim(s) 1-4, 6-10, 14-20 and 22 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-4, 6-10, 14-20 and 22 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

### Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 22 March 2004 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

### Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some \* c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- \* See the attached detailed Office action for a list of the certified copies not received.

### Attachment(s)

- 1) ☐ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☐ Information Disclosure Statement(s) (PTO/SB/08)  
Paper No(s)/Mail Date \_\_\_\_\_.
- 4) ☐ Interview Summary (PTO-413)  
Paper No(s)/Mail Date \_\_\_\_\_.
- 5) ☐ Notice of Informal Patent Application
- 6) ☐ Other: \_\_\_\_\_.

## **DETAILED ACTION**

### ***Remarks***

1. This communication is responsive to the amendment dated November 13<sup>th</sup> 2007. In the amendment dated November 13<sup>th</sup> 2007, Claims 1-4, 6-10, 14-20, and 22 are pending, Claims 1, 14, 17, and 22 are amended, Claims 5, 11-13, and 21 are canceled, and Claims 1, 14, 17, and 22 are independent Claims. The examiner acknowledges that no new matter was introduced and the amended claims are supported by the specification. This action is made FINAL.

### ***Response to Arguments***

2. Applicant's arguments dated November 13<sup>th</sup> 2007 with respect to Claims 1-4, 6-10, 14-20, and 22 have been considered but are not persuasive.

3. As to Applicant's arguments with respect to Claims 1-4, 6-10, 14-20, and 22 for the prior art(s) allegedly not teaching or suggesting "combining four samples of features into each of seven supersamples," the examiner respectfully disagrees. Pugh, col. 9, lines 29-31 with Pugh, cols. 11-12, lines 65-3 with Broder, col. 9, lines 16-22 is shown to teach this limitation claimed in the independent claims. In the cited sections, Pugh teaches 4 fingerprints and Broder teaches, essentially, any number of supersamples. For example, Broder teaches "Here we took  $k=6$ ,  $v=14$  and  $r=2$ ." "Here we took" is exemplary language meaning that a number of the values can be different from the following example. Broder, col. 9, lines 11-12 also pose a different set of parameters,

where  $k=15$ ,  $v=30$ , and  $r=4$ , and Broder, col. 8, lines 65-66 teach "the parameters can be chosen to obtain the appropriate level of filtering, i.e. the levels of acceptance and rejection." As such, the claimed "combining four samples of features into each of seven supersamples" is nothing more than a design choice of the applicant's claimed invention as taught by "the previous method." Additionally, a person of ordinary skill has good reason to pursue the known options within her or her technical grasp. If this leads to the anticipated success, it is likely the product is not of innovation but of ordinary skill and common sense.

4. Any other claims argued merely because of a dependency on a previously argued claim(s) in the arguments presented to the examiner, November 13<sup>th</sup> 2007, are moot in view of the examiner's interpretation of the claims and art and are still considered rejected based on their respective rejections from at least a prior Office action (part(s) of recited below).

### ***Response to Amendment***

#### ***Claim Objections***

5. Claims 6 and 7 are objected to because of the following informalities:

a. Claims 6 and 7 both depend on a canceled claim (Claim 5).

Appropriate correction is required.

### ***Claim Rejections - 35 USC § 112***

6. In light of the applicant's respective arguments or respective amendments, the previous 35 USC § 112 rejections to the claims have been withdrawn.

***Claim Rejections - 35 USC § 103***

7. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

8. Claims 1-4, 7-10, and 17-20 are rejected under 35 U.S.C. 103(a) as being unpatentable over U.S. Patent No. 6,349,296 (Broder et al.) in view of U.S. Patent No. 6,058,410 (Sharangpani), further in view of U.S. Patent No. 6,658,423 (Pugh et al.).

For **Claim 1**, Broder teaches: "A method for detecting similar objects in a collection of such objects, [Broder, col. 4, lines 6-15 with Broder, Fig. 3] comprising, for each of two objects:

- modifying a previous method for detecting similar objects [Broder, col. 4, lines 6-15 with Broder, Fig. 3] ...wherein the modifying comprises:
- ...each of the total number of supersamples to a number of bits of precision, [Broder, col. 9, lines 11-15] and
- requiring a number of matching supersamples out of the total number of supersamples in order to conclude that the two objects are sufficiently similar"

[Broder, col. 9, lines 1-3 with Broder, col. 9, lines 11-12 with Broder, col. 9, line 19].

Broder discloses the above limitations but does not expressly teach:

- "...so that memory requirements are reduced while avoiding false detections approximately as well as in the previous method,
- compressing...wherein the number of bits of precision is reduced from a number of bits of precision used in the previous method; and
- wherein the number of matching supersamples is greater than a number of matching supersamples required in the previous method."

With respect to Claim 1, an analogous art, Sharangpani, teaches:

- "...so that memory requirements are reduced [Sharangpani, col. 1, lines 22-27 with Broder, col. 9, lines 11-15]
- ...compressing...wherein the number of bits of precision is reduced from a number of bits of precision used in the previous method, and wherein the number of bits of precision is reduced by generating supersamples that do not include at least one least significant bit of the supersamples that were used in the previous method" [Sharangpani, col. 1, lines 22-27 with Broder, col. 9, lines 11-15].

With respect to Claim 1, an analogous art, Pugh, teaches:

- "...while avoiding false detections approximately as well as in the previous method, [Pugh, col. 3, lines 35-43]
- ... combining four samples of features into seven; [Pugh, col. 9, lines 29-31 with Pugh, cols. 11-12, lines 65-3 with Broder, col. 9, lines 16-22]

- ...wherein the number of matching supersamples is greater than a number of matching supersamples required in the previous method" [Pugh, col. 3, lines 35-43 with Broder, col. 9, lines 1-3 with Broder, col. 9, lines 11-12 with Broder, col. 9, line 19].

It would have been obvious to one of ordinary skill in the art at the time of invention having the teachings of Sharangpani and Pugh and Broder before him/her to combine Sharangpani and Pugh with Broder because the inventions are in the field of applicant's endeavor or are reasonably pertinent to the particular problem with which the applicant is concerned.

Sharangpani and Pugh's invention would have been expected to successfully work well with Broder's invention because the inventions use computers and signatures/fingerprints with bits to detect duplicates. Broder discloses a (previous) method for clustering closely resembling data objects comprising samples, supersamples, and finding similar documents. However, Broder does not explicitly disclose a reduction in samples to form a supersample, reduction in bits of precision for the fingerprints, and a greater number of matching supersamples to have objects sufficiently similar. Sharangpani discloses a method and apparatus for selecting a rounding mode for a numeric operation comprising truncating (removing) any number of bits to a desired precision. Pugh discloses detecting duplicate and near-duplicate files comprising detecting duplicates using, essentially, any number of matching fingerprints where fingerprints are combined from, essentially, any number of samples.

It would have been obvious to one of ordinary skill in the art at the time of invention having the teachings of Sharangpani and Pugh and Broder before him/her to take the removal/truncation of bits from Sharangpani, and the content of the fingerprints and matching requirements from Pugh and install them into the invention of Broder, thereby offering the obvious advantage of a reduced memory footprint (by using smaller (truncated) fingerprints/signatures) and having an reduced number of false positives.

Furthermore, it appears that the Applicant's claimed invention is a mere modification of numbers, parameters, and thresholds from the previous method. For instance, Broder, at the very least, teaches that other ranges of numbers, variables, parameters, and thresholds can be used in stating that certain numbers, variables, parameters, and thresholds were selected on an exemplary basis (Broder, col. 8, lines 62-67). As such, MPEP 2144.05 should be observed since the claimed invention appears that it is claiming an obvious optimization of ranges. Court cases of interest dealing with this are *In re Aller*, 220 F.2d 454, 456, 105 USPQ 233, 235 (CCPA 1955), *Peterson*, 315 F.3d at 1330, 65 USPQ2d at 1382, *In re Hoeschele*, 406 F.2d 1403, 160 USPQ 809 (CCPA 1969), *Merck & Co. Inc. v. Biocraft Laboratories Inc.*, 874 F.2d 804, 10 USPQ2d 1843 (Fed. Cir.), *cert. denied*, 493 U.S. 975 (1989), *In re Kulling*, 897 F.2d 1147, 14 USPQ2d 1056 (Fed. Cir. 1990), *In re Geisler*, 116 F.3d 1465, 43 USPQ2d 1362 (Fed. Cir. 1997), *In re Antonie*, 559 F.2d 618, 195 USPQ 6 (CCPA 1977), and *In re Boesch*, 617 F.2d 272, 205 USPQ 215 (CCPA 1980).

**Claim 2** can be mapped to Broder (as modified by Sharangpani and Pugh) as follows: "The method of claim 1 wherein requiring the number of matching



supersamples comprises requiring all but one of the total number of supersamples to match" [Pugh, col. 3, lines 35-43 with Broder, col. 9, lines 1-3 with Broder, col. 9, lines 11-12 with Broder, col. 9, line 19].

**Claim 3** can be mapped to Broder (as modified by Sharangpani and Pugh) as follows: "The method of claim 1 wherein requiring the number of matching supersamples comprises requiring all but two of the total number of supersamples to match" [Pugh, col. 3, lines 35-43 with Broder, col. 9, lines 1-3 with Broder, col. 9, lines 11-12 with Broder, col. 9, line 19].

**Claim 4** can be mapped to Broder (as modified by Sharangpani and Pugh) as follows: "The method of claim 1 wherein requiring the number of matching supersamples comprises requiring all supersamples to match" [Pugh, col. 3, lines 35-43 with Broder, col. 9, lines 1-3 with Broder, col. 9, lines 11-12 with Broder, col. 9, line 19].

**Claim 7** can be mapped to Broder (as modified by Sharangpani and Pugh) as follows: "The method of claim 5 wherein requiring the number of matching supersamples comprises requiring five supersamples of seven to match, [Pugh, col. 3, lines 35-43 with Pugh, cols. 11-12, lines 65-3 with Broder, col. 8, lines 62-67 with Broder, col. 9, lines 11-20] wherein the number of matching supersamples required in the previous method is two supersamples of six" [Broder, col. 9, lines 15-20].

**Claim 8** can be mapped to Broder (as modified by Sharangpani and Pugh) as follows: "The method of claim 1 wherein the objects are documents, [Broder, col. 11, lines 8-11 with Broder, col. 11, lines 19-28] and the method is used in association with a

search engine query service to determine clusters of query results that are near-duplicate documents" [Broder, col. 11, lines 8-11 with Broder, col. 11, lines 19-28].

**Claim 9** can be mapped to Broder (as modified by Sharangpani and Pugh) as follows: "The method of claim 8, further comprising selecting a single document in each cluster to report" [Pugh, col. 10, lines 50-57 or Broder, col. 10, lines 15-18].

**Claim 10** can be mapped to Broder (as modified by Sharangpani and Pugh) as follows: "The method of claim 9 wherein selecting the single document is by way of a ranking function" [Pugh, col. 10, lines 50-57].

**Claims 17-20** encompass substantially the same scope of the invention as that of Claims 1-4, respectfully, in addition to a computer-readable medium and some instructions for performing the method steps of Claims 1-4, respectfully. Therefore, Claims 17-20 are rejected for the same reasons as stated above with respect to Claims 1-4, respectfully.

9. Claim 6 is rejected under 35 U.S.C. 103(a) as being unpatentable over U.S. Patent No. 6,349,296 (Broder et al.) in view of U.S. Patent No. 6,058,410 (Sharangpani), in view of U.S. Patent No. 5,721,788 (Powell et al.), further in view of U.S. Patent No. 6,658,423 (Pugh et al.).

For **Claim 6**, Broder (as modified by Sharangpani and Pugh) teaches: "The method of claim 5 wherein:

- ...wherein the second number of bits of precision used in the previous method is 64; [Broder, col. 9, lines 11-15] and

- requiring the number of matching supersamples comprises requiring four supersamples of six to match, [Pugh, col. 3, lines 35-43 with Broder, col. 9, lines 11-20] wherein the number of matching supersamples required in the previous method is two supersamples of six” [Broder, col. 9, lines 15-20].

Broder (as modified by Sharangpani and Pugh) discloses the above limitations but does not expressly teach:

- “...compressing each supersample to the first number of bits of precision comprises recording each supersample to 16 bits of precision.”

With respect to Claim 6, an analogous art, Powell, teaches:

- “...compressing each supersample to the first number of bits of precision comprises recording each supersample to 16 bits of precision” [Powell, col. 3, lines 35-48 with Sharangpani, col. 1, lines 22-27 with Broder, col. 9, lines 11-15].

It would have been obvious to one of ordinary skill in the art at the time of invention having the teachings of Powell and Broder (as modified by Sharangpani and Pugh) before him/her to combine Powell with Broder (as modified by Sharangpani and Pugh) because both inventions are directed towards computing bits in a computer and are in the field of applicant's endeavor or are reasonably pertinent to the particular problem with which the applicant is concerned.

Powell's invention would have been expected to successfully work well with Broder (as modified by Sharangpani and Pugh)'s invention because both inventions use computers computing bits. Broder (as modified by Sharangpani and Pugh) discloses a fingerprint comprising 64-bits representing a fingerprint. However, Broder (as modified

by Sharangpani and Pugh) does not expressly disclose using a 16-bit fingerprint to represent a fingerprint/supersample. Powell discloses a method and system for digital image signatures comprising reduced (16) bits of precision for a fingerprint/signature.

It would have been obvious to one of ordinary skill in the art at the time of invention having the teachings of Powell and Broder (as modified by Sharangpani and Pugh) before him/her to take the size of the fingerprints/signatures from Powell and install it into the invention of Broder (as modified by Sharangpani and Pugh), thereby offering the obvious advantage of a reduced memory footprint (by using smaller fingerprints/signatures) and having an reduced number of false positives.

10. Claims 14-16 and 22 are rejected under 35 U.S.C. 103(a) as being unpatentable over U.S. Patent No. 6,349,296 (Broder et al.) in view of U.S. Patent No. 5,721,788 (Powell et al.), further in view of U.S. Patent No. 6,658,423 (Pugh et al.).

For **Claim 14**, Broder teaches: "A method for determining groups of near-duplicate items [Broder, col. 4, lines 6-15 with Broder, Fig. 3] in a search engine query result, [Broder, col. 11, lines 8-11 with Broder, col. 11, lines 19-28] comprising, for each of two items being compared."

Broder discloses the above limitation but does not expressly teach:

- "...combining four samples of features into each of seven supersamples;
- compressing each supersample to 16 bits of precision; and
- requiring five of the seven supersamples to match."

With respect to Claim 14, an analogous art, Pugh, teaches:

- "...combining four samples of features into each of seven supersamples; [Pugh, col. 9, lines 29-31 with Pugh, cols. 11-12, lines 65-3 with Broder, col. 9, lines 16-22]
- ...requiring five of the seven supersamples to match" [Pugh, col. 3, lines 35-43 with Broder, col. 9, lines 11-20].

With respect to Claim 14, an analogous art, Powell, teaches:

- "...compressing each supersample to 16 bits of precision" [Powell, col. 3, lines 35-48].

It would have been obvious to one of ordinary skill in the art at the time of invention having the teachings of Powell, Pugh and Broder before him/her to combine Powell and Pugh with Broder because the inventions are in the field of applicant's endeavor or are reasonably pertinent to the particular problem with which the applicant is concerned.

Powell's and Pugh's inventions would have been expected to successfully work well with Broder's invention because the inventions use computers and signatures/fingerprints with bits to detect duplicates. Broder discloses a (previous) method for clustering closely resembling data objects comprising samples, supersamples, and finding similar documents. However, Broder does not explicitly disclose a reduction in samples to form a supersample, reduction in bits of precision for the fingerprints, and a greater number of matching supersamples to have objects sufficiently similar. Powell discloses a method and system for digital image signatures comprising reduced (16) bits of precision for a fingerprint. Pugh discloses detecting

duplicate and near-duplicate files comprising detecting duplicates using, essentially, any number of matching fingerprints where fingerprints are combined from, essentially, any number of samples.

It would have been obvious to one of ordinary skill in the art at the time of invention having the teachings of Powell, Pugh and Broder before him/her to take the size of the fingerprints/signatures from Powell, and the content of the fingerprints and matching requirements from Pugh and install them into the invention of Broder, thereby offering the obvious advantage of a reduced memory footprint (by using smaller fingerprints/signatures) and having an reduced number of false positives.

Furthermore, it appears that the Applicant's claimed invention is a mere modification of numbers, parameters, and thresholds from Broder's method. For instance, Broder, at the very least, teaches that other ranges of numbers, variables, parameters, and thresholds can be used in stating that certain numbers, variables, parameters, and thresholds were selected on an exemplary basis (Broder, col. 8, lines 62-67). As such, MPEP 2144.05 should be observed since the claimed invention appears that it is claiming an obvious optimization of ranges. Court cases of interest regarding this are *In re Aller*, 220 F.2d 454, 456, 105 USPQ 233, 235 (CCPA 1955), *Peterson*, 315 F.3d at 1330, 65 USPQ2d at 1382, *In re Hoeschele*, 406 F.2d 1403, 160 USPQ 809 (CCPA 1969), *Merck & Co. Inc. v. Biocraft Laboratories Inc.*, 874 F.2d 804, 10 USPQ2d 1843 (Fed. Cir.), *cert. denied*, 493 U.S. 975 (1989), *In re Kulling*, 897 F.2d 1147, 14 USPQ2d 1056 (Fed. Cir. 1990), *In re Geisler*, 116 F.3d 1465, 43 USPQ2d

1362 (Fed. Cir. 1997), *In re Antonie*, 559 F.2d 618, 195 USPQ 6 (CCPA 1977), and *In re Boesch*, 617 F.2d 272, 205 USPQ 215 (CCPA 1980).

**Claim 15** can be mapped to Broder (as modified by Powell and Pugh) as follows:  
“The method of claim 14, further comprising selecting a single document in each cluster to report” [Pugh, col. 10, lines 50-57 or Broder, col. 10, lines 15-18].

**Claim 16** can be mapped to Broder (as modified by Powell and Pugh) as follows:  
“The method of Claim 15 wherein selecting the single document is by way of a ranking function” [Pugh, col. 10, lines 50-57].

**Claim 22** encompasses substantially the same scope of the invention as that of Claim 14, in addition to a computer-readable medium and some instructions for performing the method steps of Claim 14. Therefore, Claim 22 is rejected for the same reasons as stated above with respect to Claim 14.

11. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire **THREE MONTHS** from the mailing date of this action. In the event a first reply is filed within **TWO MONTHS** of the mailing date of this final action and the advisory action is not mailed until after the end of the **THREE-MONTH** shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of

Application/Control Number:  
10/805,805  
Art Unit: 2161

Page 15

the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.



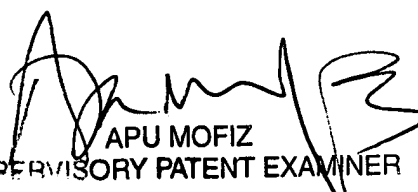
**Conclusion**

12. Any prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Applicant is advised that, although not used in the rejections above, prior art cited on any PTO-892 form and not relied upon is considered materially relevant to the applicant's claimed invention and/or portions of the claimed invention.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Brent S. Stace whose telephone number is 571-272-8372 and fax number is 571-273-8372. The examiner can normally be reached on M-F 9am-5:30pm EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Apu M. Mofiz can be reached on 571-272-4080. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300. Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Brent Stace *B.S.*

  
APU MOFIZ  
SUPERVISORY PATENT EXAMINER